ELSEVIER

FEMS Microbiology Ecology 44 (2003) 231–242

FEMS MICROBIOLOGY Ecology

www.fems-microbiology.org

Molecular analysis of bacterial microbiota in the gut of the termite *Reticulitermes speratus* (Isoptera; Rhinotermitidae)

Yuichi Hongoh [a,b,*], Moriya Ohkuma [a,b], Toshiaki Kudo [a,b,c]

[a] *International Cooperative Research Project, Japan Science and Technology Corporation (JST-ICORP), Saitama, Japan*

[b] *Molecular Microbial Ecology Division, Bioscience Technology Center, RIKEN (The Institute of Physical and Chemical Research), Wako, Saitama 351-0198, Japan*

[c] *Division of Environmental Molecular Biology, Graduate School of Integrated Science, Yokohama City University, Kanagawa 230-0045, Japan*

Received 5 August 2002; received in revised form 10 November 2002; accepted 29 December 2002

First published online 31 January 2003

Abstract

The molecular diversity and community structure of bacteria from the gut of the termite *Reticulitermes speratus* were analyzed by the sequencing of near-full-length 16S rRNA genes, amplified by polymerase chain reaction. The results of the analysis of 1344 clones indicated a predominance of spirochetes in the gut. Spirochetal clones accounted for approximately half of the analyzed clones. The clones related to *Bacteroides*, Clostridia, and the candidate division Termite Group I each accounted for approximately 5–15% of the analyzed clones. The rest were comprised of Proteobacteria, Actinobacteria, *Mycoplasma* and others. Using the criterion of 97% sequence identity, the clones were sorted into 268 phylotypes, including 100 clostridial, 61 spirochetal and 31 *Bacteroides*-related phylotypes. More than 90% of the phylotypes were found for the first time, and some constituted monophyletic clusters with sequences recovered from the gut of other termite species.

Keywords: Diversity; Termite; Symbiosis; Gut bacterium

1. Introduction

Termites harbor an abundance and diversity of gut microbes, comprised of bacteria, protists, fungi and archaea. The termite gut microbe ecosystem is one of the most fascinating examples of symbiosis between an animal and microbes, and among a diversity of microbes. The characteristics of these gut microbes have been extensively studied, and it is believed that the microbes provide carbon, nitrogen and energy nutrition to their host termites, to the extent that termites can no longer live without them [1–3]. However, difficulties in the cultivation of these gut microbes and the complexity of their community structure have been obstacles to a comprehensive understanding of the symbiotic system. To overcome these problems, efforts to cultivate and characterize as yet unculturable microbes have continued [4,5] and culture-independent approaches have been applied [6–11]. In our previous work, partial 16S rRNA genes were amplified by polymerase chain reaction (PCR) from the gut of the termite *Reticulitermes speratus* to estimate the diversity of the gut bacterial microbiota [6]. We found a predominance of five bacterial groups: the *Cytophaga-Flexibacter-Bacteroides* (CFB) group; the low-G+C Gram-positive bacteria; Proteobacteria; Spirochaeta; and the newly discovered 'Termite Group I' (TG I). However, we analyzed only 55 clones and sequence lengths of approximately 300 bp in that study, insufficient for the reliable elucidation of bacterial microbiota and their phylogenetic affiliation. While the investigation of microbiota using PCR amplification and the sequencing of small subunit rRNA genes contains some problems, such as PCR errors, PCR bias and different copy numbers of the gene, it is widely accepted as a powerful and essential tool for the investigation of as yet unculturable microbial communities. In studies of termite gut microbe symbiosis, thorough molecularly based investigations of microbiota are needed not only for studies of their phylogenetic diversity and evolution, but also for

* Corresponding author. Tel.: +81 (48) 467-9546; Fax: +81 (48) 462-4672.

E-mail address: yhongo@postman.riken.go.jp (Y. Hongoh).

0168-6496/03/$22.00
doi:10.1016/S0168-6496(03)00026-6

evaluating the ecological and physiological impact of various gut microbes on a termite's metabolism.

In the present paper, we report a detailed analysis of bacterial diversity and community structure in the gut of the termite *R. speratus*. Data are based on sequences of near-full-length 16S rRNA genes, amplified by PCR. To reduce the possibility of overlooking any major bacterial populations in the analysis, several sets of Bacteria-universal primers and annealing temperatures were used.

2. Materials and methods

2.1. Sample collection and extraction of DNA

Wood-feeding lower termites, *R. speratus* (family Rhinotermitidae), were collected at Ogose, Saitama, Japan. Within 3 h of collection, the termites were washed in sterilized water and their guts were drawn out using sterilized forceps. The isolated guts of 30 individuals were submerged in solution U [11] on ice and gently crushed using a pestle. After centrifugation at 2000×*g* for 5 min, the precipitate was treated with 0.2 mg ml^{-1} Zymolyase 100-T (Seikagaku) and 2 mg ml^{-1} lysozyme (Nacalai) at 30°C for 30 min. The reaction mixture was dropped into liquid nitrogen and crushed into a fine powder, from which DNA was extracted using the Qiagen RNA/DNA kit.

2.2. PCR amplification and preparation of clone libraries

16S rRNA genes were amplified from the gut sample by PCR, using Bacteria-universal primers (Table 1). PCR was performed using a PTC-200 Thermal Cycler (MJ Research), *EX-Taq* polymerase (Takara), and the following program: 2 min of initial denaturation at 95°C followed by 12 or 18 cycles of denaturation (30 s at 95°C), annealing (60 s at 45, 50 or 55°C), and extension (4 min at 72°C), with a final extension at 72°C for 10 min. The concentrations of the template and primers were 0.1 ng μl^{-1} and 1 μM, respectively. Fourteen sets of PCRs, using different primers and the above annealing temperatures and cycle numbers, were performed to prepare independent clone libraries (Fig. 1). In each PCR set, the products were purified from 200 μl of reaction mixture, using the QIAquick PCR Purification Kit (Qiagen), followed by ethanol precipitation. The products were cloned into pCR-4 TOPO TA cloning vectors (Invitrogen), and clone libraries were established on LB agar plates.

Table 1
Bacteria-universal primers for PCR and sequencing

Primer	Sequence	Reference
27F	5′-AGAGTTTGATCMTGGCTCAG-3′	[38]
39F	5′-TGGCTCAGRWYGAACGCTRG-3′	This study
41F	5′-GCTCAGATTGAACGCTGGCG-3′	[39]
63F	5′-CAGGCCTAACACATGCAAGTC-3′	[40]
64F	5′-BGYCTWANRCATGCAAGTYG-3′	This study
533F	5′-TGCCAGCAGCCGCGGTAA-3′	[41][a]
908R	5′-CGTCAATTCMTTTGAGTT-3′	[41]
1389R	5′-ACGGGCGGTGTGTACAAG-3′	[40,42]
1492R	5′-GGTTACCTTGTTACGACTT-3′	[38]

[a]Slightly modified from the original primer.

2.3. Sequencing and phylogenetic affiliation of clones

Ninety-six clones were randomly chosen from each clone library for analysis. Approximately 550 bp of the 16S rDNA clones, corresponding to the region 563–1114 in *Escherichia coli* (J01695), were sequenced with primer 533F (Table 1), using the Big-Dye Terminator Cycle Sequencing Kit (Perkin-Elmer) and ABI 3700 or 377 genetic analyzers.

The clones sharing more than 98% sequence identity were grouped into the same phylotype, and the full length of representative clones of all the phylotypes was sequenced with primers T7, T3 and 908R (Table 1), using the long read protocol distributed by PE Biosystems. They were further sorted into phylotypes using the criterion of 97% sequence identity. All the sequences were subjected to the Check-Chimera program on the Ribosomal Database Project (RDP) web site [12] for the elimination of chimeric sequences, and analyzed by BLAST 2.0 [13] and FASTA 3.0 [14] to find the closest phylogenetic neighbors. The sequences of all the phylotypes were aligned with those which marked the highest point in the FASTA analysis and with other representative sequences from on-line databases, using CLUSTAL X1.8 [15]. The region corresponding to positions 64–1388 in *E. coli* (J01695) was used for the alignment, and all the sites containing gaps were excluded from the phylogenetic analysis to decrease ambiguity. As for shorter sequences, only those which were thought to be essential were included, and the sites containing missing data were excluded pairwise in the phylogenetic analysis. A neighbor-joining tree of the aligned sequences was constructed using MEGA V2.1 [16], based on the distance matrix calculated according to the Jukes–Cantor model. The tree was tested statistically using a bootstrap resampling of 1000 times. GENETYX V10.1.6 (Software Development) was used to calculate sequence similarities. DNASIS V3.7 (Hitachi) was used for the general manipulation of the sequence data. The sequence data will appear in the DDBJ, EMBL and GenBank nucleotide sequence databases under accession numbers AB088857–AB089126.

2.4. Statistical analysis

The difference between clone libraries was examined pairwise by the statistical analysis described by Singleton et al. [17]. From each clone library, 96 sequences of approximately 550 bp, corresponding to the region 563–1114 in *E. coli* (J01695), were used for inter-clone comparisons.

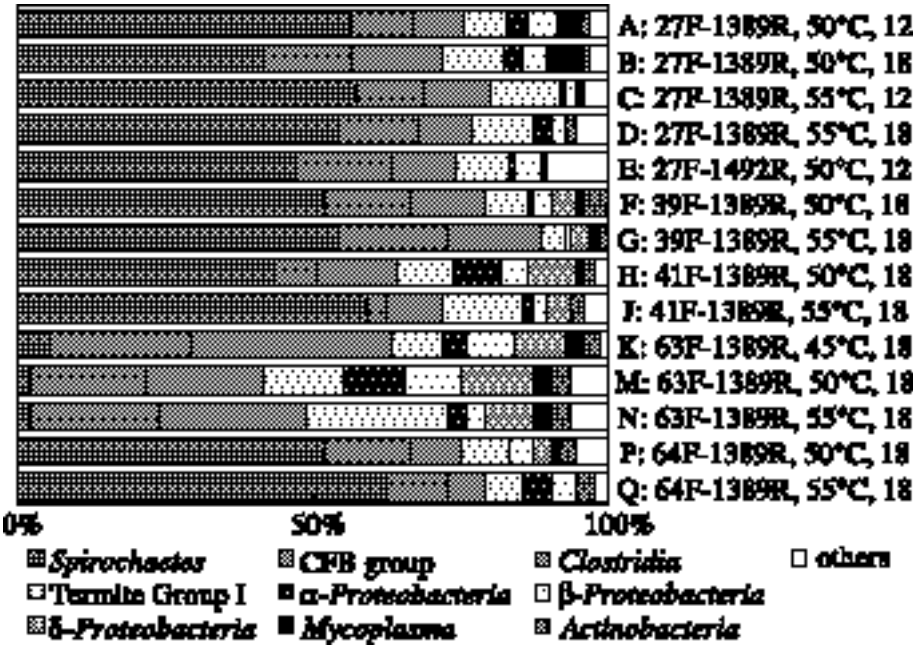


Fig. 1. Relative abundance of 16S rDNA clones in the 14 clone libraries A–Q. The primer sets, annealing temperatures and number of cycles for PCR are attached to the library names. Ninety-six clones were analyzed for each clone library. The clones were affiliated with the taxonomic groups by phylogenetic analysis. Rare taxa were grouped as 'others'.

Distance matrices were calculated using MEGA V2.1, according to the Jukes–Cantor model and based on the sequences aligned by CLUSTAL X1.8, and were analyzed using the program LIBSHUFF [17], distributed freely at http://www.arches.uga.edu/~whitman/libshuff.html. The coverage [18] was calculated by the formula $[1-(n/N)]$, where n is the number of phylotypes represented by only one clone and N is the total number of clones.

3. Results and discussion

3.1. Estimation of bacterial diversity and community structure in the termite gut

From the total of 1344 (96×14) clones, 268 phylotypes were recognized. The phylotypes were classified phylogenetically into 11 bacterial divisions, an unidentified cluster and an unidentified sequence (Table 2). The relative abundance of clones affiliated with the respective taxonomic groups (Table 2) is shown in Fig. 1. No statistically significant difference between the clone libraries was detected using the program LIBSHUFF [17], except in the comparison between a library prepared from PCR using primer 63F and a library from PCR using the other primers ($P<0.01$). The difference was obviously due to the lack of spirochetal clones in libraries K, M and N (Fig. 1), presumably caused by preferential annealing due to six or seven mismatched base pairs between 63F and the corresponding region of the spirochetal sequences. These three clone libraries were not used in the estimation of community structure and diversity.

In all the clone libraries, except for K, M and N, spirochetal clones were the most dominant, accounting for 42–63% of the sequenced clones (Fig. 1). This suggests a predominance of spirochetes in the gut of *R. speratus*, which is comparable to the result of the microscopic observation of the gut contents of the higher termite *Nasutitermes lujae* [19]. The *Bacteroides*-related, clostridial and TG I clones formed the second most dominant groups, accounting for 3–19%, 6–16%, and 4–11% of the sequenced clones, respectively.

The community structure at the phylotype level was outlined in the species abundance curve depicted in Fig. 2.

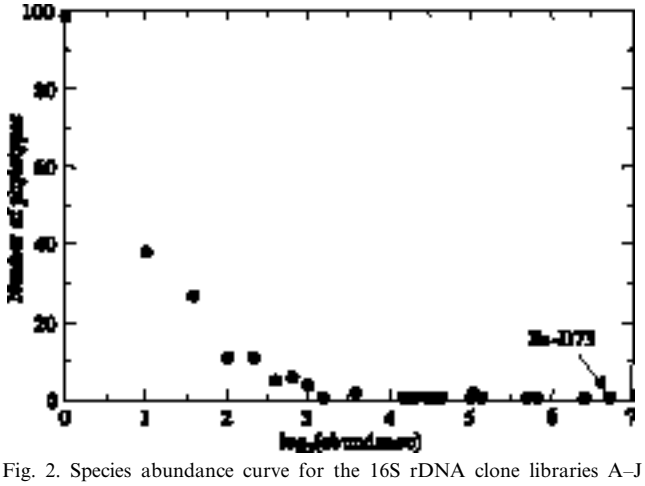


Fig. 2. Species abundance curve for the 16S rDNA clone libraries A–J and P–Q. The data from the clone libraries K, M and N are excluded because of their lack of the most dominant, spirochetal clones. The total numbers of clones and phylotypes were 1056 and 218, respectively. The 16S rDNA phylotype Rs-D73 is indicated as the most abundant.

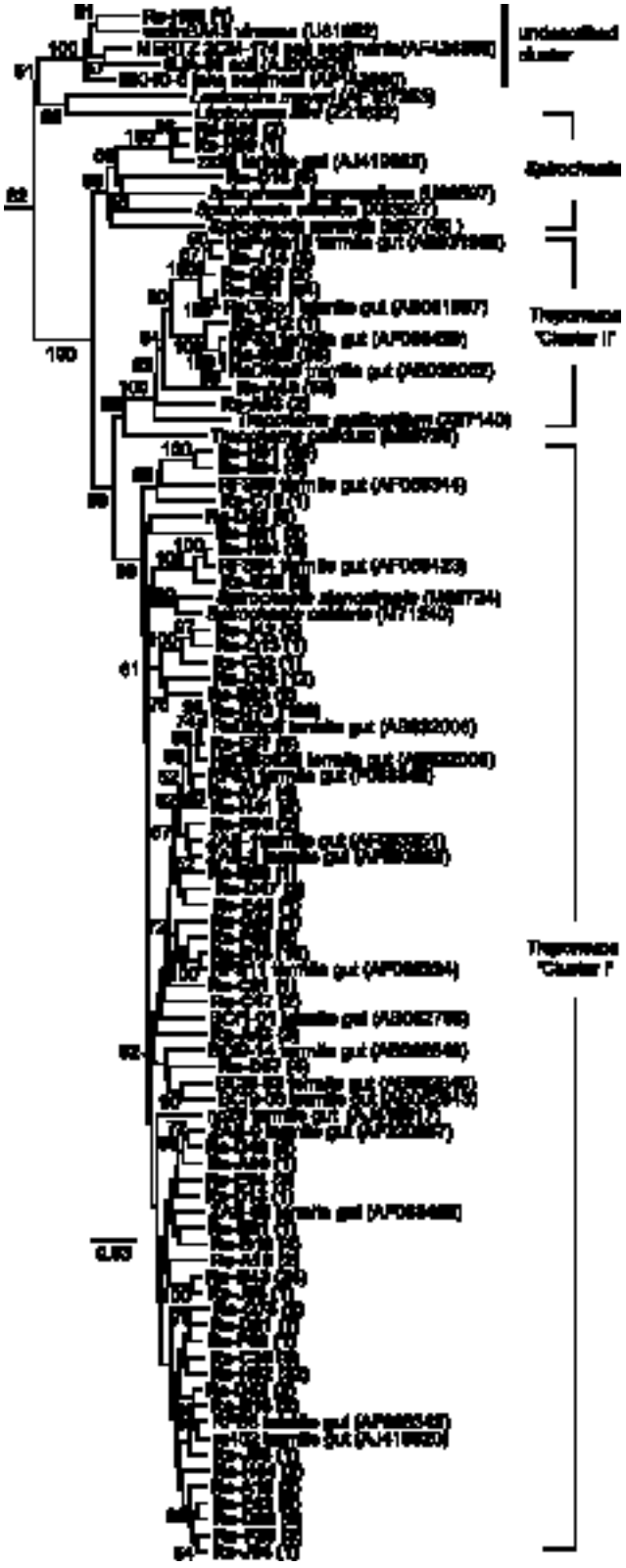
Table 2
Classification and number of 16S rDNA phylotypes found in this study

Taxon	RDP code	Number of phylotypes[a]	Accession number
Spirochaeta			
Treponema	[2.27.3.2]	57	AB088857–916
Spirochaeta	[2.27.3.1,3]	3	
Undescribed cluster	none[b]	1	AB089126
CFB group			
Bacteroides and relatives	[2.15.1]	31	AB088917–47
Low-G+C Gram-positives			
Clostridium/*Eubacterium*	[2.30.4, 5, 9]	100	AB088948–9047
Mycoplasma	[2.30.8.4]	7	AB089053–60
Lactococcus, *Enterococcus*	[2.30.7.20, 21]	3	AB089061–3
Anaerobaculum	[2.11]	3	AB089064–6
Termite Group I	[2.19]	5	AB089048–52
Actinobacteria	[2.30.1]	14	AB089069–82
Proteobacteria			
α	[2.28.1]	16	AB089083–99
β	[2.28.2]	4	AB089100–3
δ	[2.28.4.1, 10]	6	AB089104–10
ε	[2.28.5]	3	AB089111–3
γ	[2.28.3]	2	AB089115–6
Planctomycetes	[2.20]	4	AB089117–20
Verrucomicrobia	[2.10]	2	AB089121–22
Cyanobacteria	[2.21]	1	AB089123
Acidobacteria	[2.25.3.6]	1	AB089124
OP11	[2.28.4.2]	2	AB089114, -125
Unidentified cluster	none	2	AB089067–8
Unidentified bacterium	none	1	AB089107
Total number of phylotypes		268	

[a]Phylotypes were defined using the criterion of 97% sequence identity in the region corresponding to positions 64–1388 in *E. coli* (J01695).
[b]The reference sequences from the databases in this cluster are classified into different RDP taxonomic groups as follows. SJA-88: *Nitrospina* subdivision [2.13.6]; vadinBA43: *Heliobacterium* group [2.30.3.2] in the low-G+C Gram-positive bacteria; EKHO-6; *Thermoanaerobacter* and relatives [2.30.2] in the low-G+C Gram-positive bacteria. Also see Fig. 3.

Whereas a species abundance of prokaryotes is theoretically deduced to approximate a log-normal distribution [20], namely, few numerically dominant and few very rare species but many species of intermediate abundance, as demonstrated in the gut bacteria of pigs [21], it was quite different in the present study. We found few dominant phylotypes and many rare phylotypes (Fig. 2). The number of sequenced clones in the study of gut bacteria from pigs [21] was up to 4270, much more than the 1056 (without the clones from libraries K, M and N) in our study and the coverage, 0.978 [21], was also much higher than 0.905 in the present study, i.e., 2.2% of the additionally sequenced clones from pig intestine would represent new phylotypes while 9.5% of those in the present study would. These differences imply that the number of sequenced clones in the present study is insufficient to reflect the diversity of all bacteria in the termite gut. However, Curtis et al. [20] recently proposed a mathematical equation to estimate prokaryotic diversity in an environment from a practically acquirable data set of 16S rDNA clone libraries. To solve the formula [20], we need the values N_T, N_{max} and N_{min}, where N_T is the total number of individuals in the community, N_{max} is the number of individuals in the most abundant species and N_{min} is the number of individuals in the least abundant species. N_T can be estimated from a previous study [22] using *Reticulitermes flavipes*, a congeneric of *R. speratus* with similar body size. In *R. flavipes*, the total count of prokaryotic cells was around 10^{10} cells ml^{-1} gut fluid, corresponding to approximately 3×10^6 cells gut^{-1} [22]. The population of archaeal cells can be ignored because they were estimated to occupy less than 10% of the prokaryotic community [10]. N_{max} can be estimated from N_T and the frequency of the most abundant phylotype, Rs-D73, which comprised approximately 10% of the sequenced clones from all the clone libraries, except for K, M and N. Assuming N_{min} is 1, as suggested by the authors [20], the equation was solved for the number of phylotypes to be about 6000 phylotypes ml^{-1} gut fluid and 700 phylotypes gut^{-1}. However, we should be careful in handling this type of approximation. Besides the intrinsic problems of this estimation, especially when a 16S rDNA clone library is used as the data [23], using the frequency of Rs-D73 as the most abundant phylotype could be problematic. Since this phylotype was identical to one of the dominant bacterial phylotypes amplified by PCR from the gut protist *Dinenympha porteri* [11], it is possible that it is one of the protist ectosymbionts and that its habitat is restricted to the cell surface of

D. porteri. In this case, the abundance of Rs-D73 might not be appropriate for estimating diversity in the whole gut. To more accurately estimate the microbial diversity in termite guts, further ecological studies, including the localization and characterization of dominant bacteria, are needed.



3.2. Phylogenetic affiliation of 16S rDNA phylotypes

3.2.1. Spirochaeta

As described above, spirochetes comprised approximately half of the sequenced clones. Sixty-one phylotypes were affiliated with this group and 51 of them were found for the first time. Most of the phylotypes (57 of 61) were affiliated with the genus *Treponema* (Table 2 and Fig. 3). They were further classified into two distinct clusters, 'termite *Treponema* clusters I and II', as designated in our previous studies [11,24]. Fifty phylotypes were affiliated with cluster I and seven with cluster II. The sequence similarities between the 50 phylotypes in cluster I were more than 87% and those of the seven phylotypes in cluster II were more than 88%. The most abundant phylotype in all of the sequenced clones, Rs-D73, was included in cluster I and was almost identical (>99%) to RsDiSp8 (AB032006), which was recovered from the protist *D. porteri* in the gut of *R. speratus* [11]. The second most abundant phylotype in all of the sequenced clones, Rs-G65, was affiliated with cluster II and almost identical (>99%) to RsDiSp9 (AB032002), also recovered from *D. porteri* [11]. The phylotype Rs-E01, sharing 98.7% sequence identity with the *D. porteri*-derived sequence RsDiSp1 (AB031997), was also abundantly found in the sequenced clones (Figs. 3 and 4). These phylotypes implied an abundance of spirochetal ectosymbionts with the protists in the gut of *R. speratus*. In addition to these three phylotypes, several abundant phylotypes were included in the *Treponema* clusters as shown in Fig. 4.

Although the physiological characteristics of bacteria cannot be deduced from their phylogenetic position alone, the predominance of treponemes in the gut of *R. speratus*, as suggested in this study, implies interesting possibilities. The *Treponema* spp. in cluster I, ZAS-1 and ZAS-2, isolated from the gut of the termite *Zootermopsis angusticollis* were characterized as H_2/CO_2 acetogens [4], and ZAS-9 was characterized as a N_2 fixer [5]. These microbial activities are thought to be essential for the nutrition of wood-feeding termites [2]. Therefore, it is worth attempting to

←
Fig. 3. Phylogenetic tree showing the positions of 16S rDNA phylotypes affiliated with Spirochaeta, recovered from the gut of the termite *R. speratus*. The total number of clones belonging to each phylotype is indicated in parentheses. Only bootstrap confidence values greater than 50 are indicated. The scale bar represents a 3% estimated sequence divergence. The 'undescribed cluster' includes the short sequences MERTZ 2CM-174 and EKHO-6 (corresponding to positions 518–1388 in *E. coli* J01695). EKHO-6 was found in the databases as a member of the candidate division Sediment-4. The origins of the reference sequences from termite guts are as follows. za29: *Zootermopsis angusticollis*; RsDiSp series: *R. speratus* [11]; HsDiSp320 and HsPySp15: *Hodotermopsis sjoestedti* [11]; RFS series: *R. flavipes* [9]; ZAS series: *Z. angusticollis* [5,9]; nc26: *Neotermes castaneus*; BCf series: *Coptotermes formosanus*.

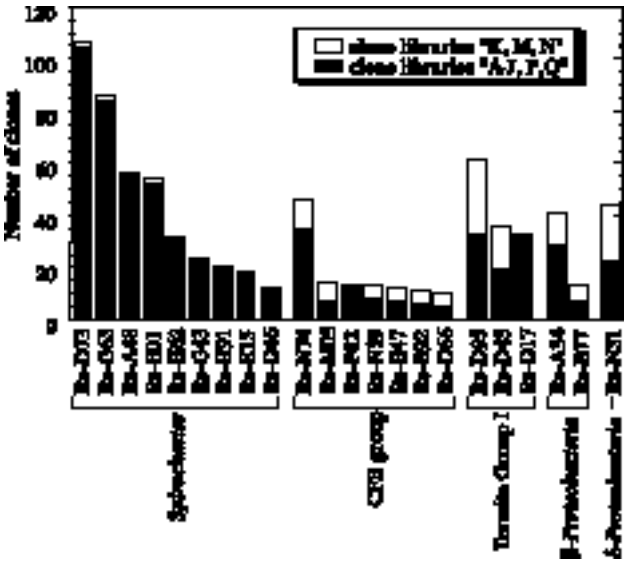


Fig. 4. Abundance of 16S rDNA phylotypes frequently found from all the sequenced clones. Since the clone libraries K, M and N lacked spirochetal clones, their data are distinguished from the others. Only phylotypes represented by more than 10 clones are shown.

characterize abundant phylotypes such as Rs-D73, Rs-G65 and Rs-A48 for these activities using culture-dependent and/or -independent methods.

Only a few spirochetal phylotypes were affiliated outside *Treponema*. Three phylotypes were affiliated with the genus *Spirochaeta* and two of the three constituted a monophyletic cluster with za29 (AJ419823), the sequence amplified from the gut of *Z. angusticollis* (Fig. 3). The phylotype Rs-H88 was clustered with four unidentified environmental samples and its taxonomic position is uncertain (Table 2 and Fig. 3).

3.2.2. Bacteroides *and relatives*

The *Bacteroides*-related clones were sorted to 31 phylotypes; 28 of them were found for the first time. The three phylotypes Rs-N74, Rs-B92 and Rs-M21 were almost identical (>99%) to the sequences UN71 (D63629), UN78 (D63632) and UN108 (AB055736), respectively, which were also recovered from the gut of *R. speratus* [6]. Since these 'UN-series' sequences were as short as about 850 bp, we excluded them from the phylogenetic analysis. One of the three, Rs-N74, was the most abundant *Bacteroides*-related phylotype from the sequenced clones (Figs. 4 and 5). It constituted a monophyletic cluster exclusively comprised of termite gut-derived clones (corresponding to cluster V in our previous study [25]), although the bootstrap confidence values for the roots were not significantly high (Fig. 5 and [25]). The sequences listed on the top of the BLAST results for 19 *Bacteroides*-related phylotypes, including the above three, were from the guts of various termites. Since all the related termite-derived sequences were as short as about 850 bp except for the BCf series from the termite *Coptotermes formosanus* (only found in databases as the clone series 'BCf'), they were not included in the present analysis. Instead, a detailed phylogenetic analysis of these shorter *Bacteroides*-related sequences, including the above UN series, was done and reported in our previous paper [25]. The sequence similarities between the 31 *Bacteroides*-related phylotypes were 78–95%.

3.2.3. The low-G+C Gram-positive bacteria

The clones affiliated with the low-G+C Gram-positive bacteria were very diverse and sorted to 113 phylotypes. All of them were found for the first time. Most of them, 100 of 113, were affiliated with Clostridia (Table 2), consisting of 44 phylotypes related to *Clostridium leptum* and *C. thermocellum* (Fig. 6), 23 phylotypes related to bovine ruminal clones, swine intestinal clones and termite gut clones in unidentified clusters (Fig. 7), 23 phylotypes related to *C. propionicum* and *C. xylanolyticum* (Fig. 8A),

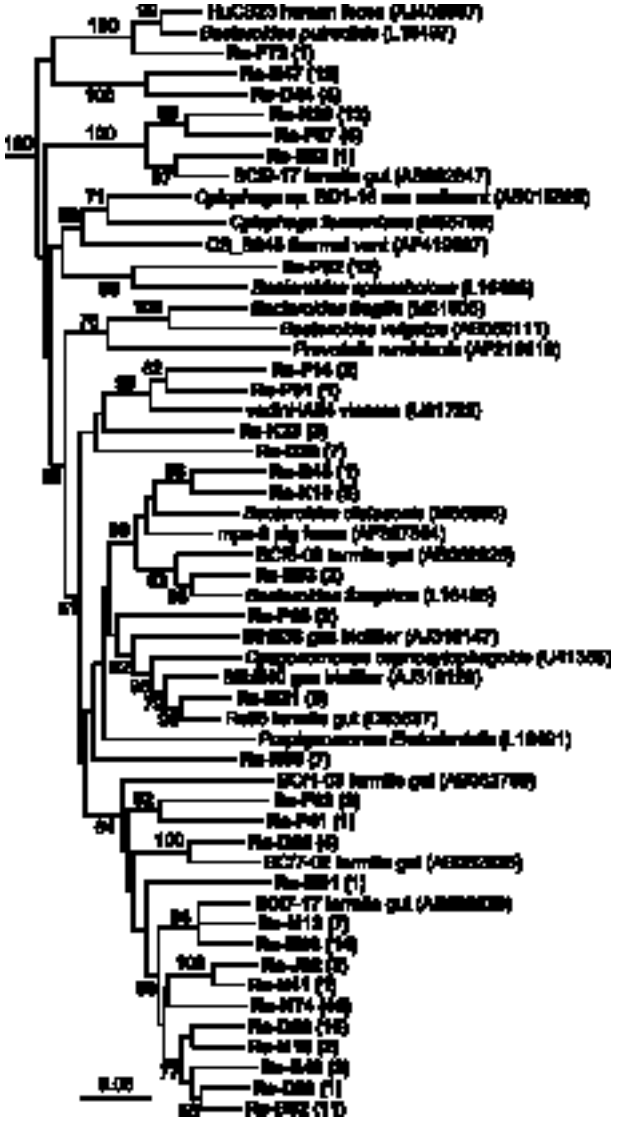


Fig. 5. Phylogenetic tree showing the positions of 16S rDNA phylotypes related to *Bacteroides* in the CFB group, recovered from the gut of the termite *R. speratus*. See the legend of Fig. 3 for explanation.

and 10 phylotypes related to *C. purinolyticum* (Fig. 8B). All of these phylotypes were found in small numbers (one to nine) from the sequenced clones. The BLAST analysis of the 100 clostridial phylotypes revealed that the database sequences showing the highest homology to 28 phylotypes

Fig. 6. Phylogenetic tree showing the positions of 16S rDNA phylotypes related to *C. leptum* and *C. thermocellum* in the low-G+C Gram-positive bacteria, recovered from the gut of the termite *R. speratus*. *Clostridium termitidis* [43] and *Termitobacter aceticus* were from the termite *N. lujae*. See the legend of Fig. 3 for explanation.

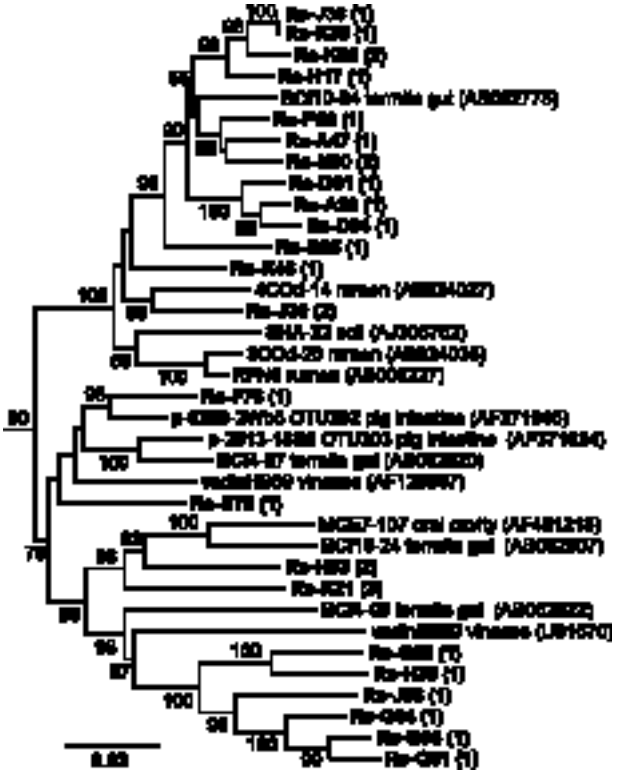


Fig. 7. Phylogenetic tree showing the positions of 16S rDNA phylotypes from the gut of the termite *R. speratus*, affiliated with clusters comprised only of sequences from uncultured clostridia in the low-G+C Gram-positive bacteria. The apparent identity between the phylotypes Rs-J36 and Rs-K36 was due to the exclusion of sites containing gaps. See the legend of Fig. 3 for explanation.

were from termite guts (the BCf series, described above), and those to 46 phylotypes were from the intestines or feces of other animals such as pigs [21], cows [26,27], chickens [28] and humans [29]. The cluster shown in Fig. 6 contained no described bacteria and corresponded to the cluster of bovine ruminal clones and Cluster-E of Leser et al. [21].

Three phylotypes, each represented by one to four clones, were affiliated with lactic acid bacterial clusters (Fig. 9A). Rs-D42 formed a monophyletic cluster with *Enterococcus*, and Rs-C68 and Rs-B70 clustered with *Lactococcus*. The closest sequence in databases to Rs-B70 was BCf2-19 (94.7% similarity, Fig. 9A) from *C. formosanus*. Seven phylotypes affiliated with *Mycoplasma* constituted a monophyletic cluster with the sequences from *C. formosanus* and a sequence amplified from a cell of *Koruga bonita*, a symbiotic protist in the gut of the termite *Mastotermes darwiniensis* [30], as shown in Fig. 9B. The sequence similarities among these termite-derived mycoplasmas were 79–95%. Three phylotypes represented by one to seven clones were affiliated with the *Anaerobaculum thermoterrenum* cluster (Fig. 9C). Two of the three formed a monophyletic cluster with the two sequences from *C. formosanus*, and their similarities were 93–96%.

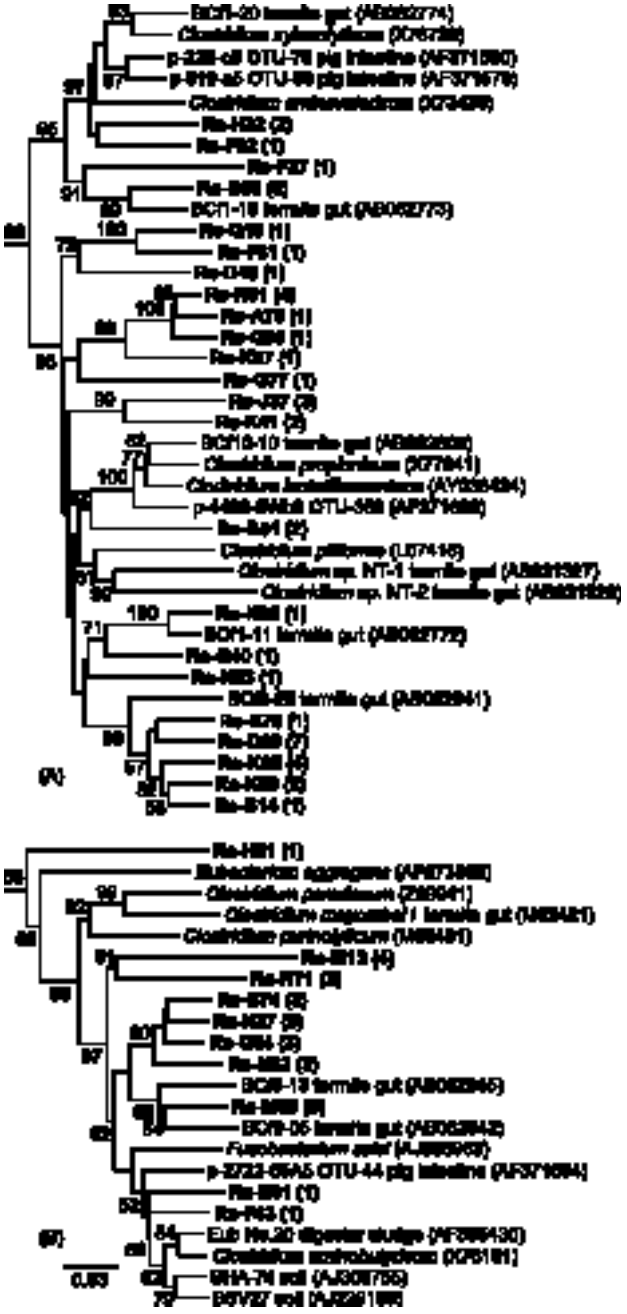


Fig. 8. Phylogenetic tree showing the positions of 16S rDNA phylotypes related to *C. propionicum* (A) and *C. purinolyticum* (B) in the low-G+C Gram-positive bacteria, recovered from the gut of the termite *R. speratus*. See the legend of Fig. 3 for explanation.

3.2.4. Termite Group I

This group of bacteria was first recognized in our previous study, as the 16S rDNA clones TG13 (AB004575) and TG14 (AB004576) amplified from the gut of *R. speratus* [6]. In the present study, five phylotypes were affiliated with this group, and the two phylotypes Rs-D95 and Rs-D17 showed high sequence similarities to TG13 (98.3%) and TG14 (99.9%), respectively (Fig. 10). As in our previous study [6], clones belonging to this group were found abundantly in the sequenced clones, forming one of the second most dominant groups (Figs. 1 and 4). The similarities between these five phylotypes were more than 95% and constituted a robust monophyletic cluster (Fig. 10). This cluster was related to four environmental samples, 4COd-3 from a bovine rumen [27], SBR2095 from activated sludge [31], vadinHB76 from wine vinasse [32], and a short sequence WCHA1-89 from a contaminated aquifer [33]. Whereas the high bootstrap confidence value (99%) supported the monophyly of the termite cluster and the four environmental samples, the sequence similarities between them were below 82%. No information, except for 16S rDNA sequences, has been reported so far on this group.

3.2.5. Proteobacteria

Thirty-one phylotypes were affiliated with Proteobacteria and further classified into five subdivisions: 16 α-Proteobacteria, two γ-Proteobacteria, four β-*Proteobacteria*, three ε-Proteobacteria, and six δ-Proteobacteria (Fig. 11). The 16 α-proteobacterial phylotypes were represented

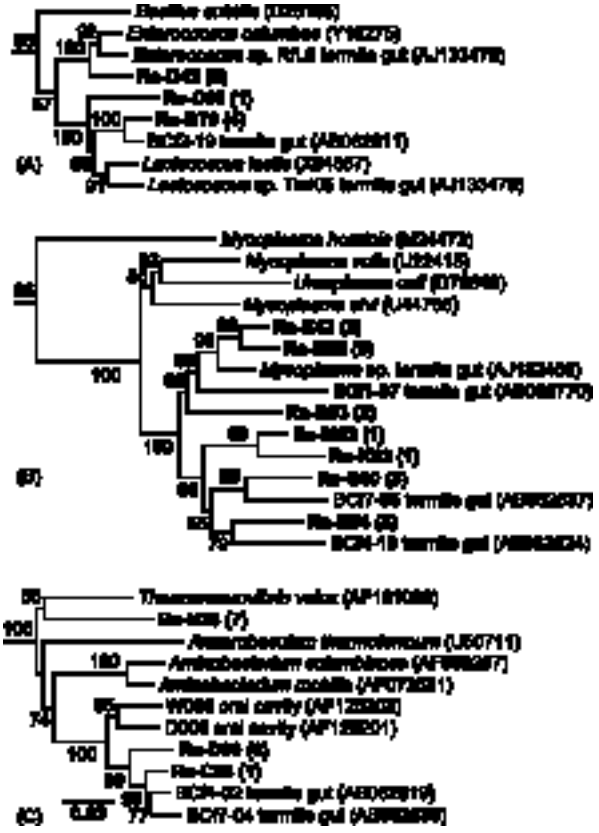


Fig. 9. Phylogenetic tree showing the positions of 16S rDNA phylotypes affiliated with Bacilli (A), *Mycoplasma* (B) and *Anaerobaculum thermoterrnum* cluster (C) in the low-G+C Gram-positive bacteria, recovered from the gut of the termite *R. speratus*. The isolates *Enterococcus* sp. RfL6 and *Lactococcus* sp. Tml05 were from the guts of the termites *R. flavipes* and *Thoracotermes macrothorax*, respectively [44]. The clone *Mycoplasma* sp. (AJ132469) is a short sequence (corresponding to positions 255–1388 in *E. coli* J01695), amplified from a single cell of the protist *Koruga bonita* in the gut of the termite *Mastotermes darwiniensis* [30]. See the legend of Fig. 3 for explanation.

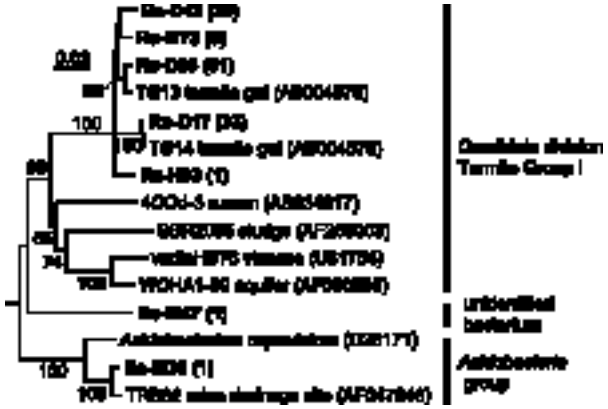


Fig. 10. Phylogenetic tree showing the positions of 16S rDNA phylotypes affiliated with the candidate division TG I and Acidobacteria group, recovered from the gut of the termite *R. speratus*. The clones TG13 and TG14 were previously recovered from *R. speratus* [6]. The clone WCHA1-89 from a contaminated aquifer [33] is a short sequence, corresponding to positions 534–1388 in *E. coli* (J01695). The phylotype Rs-M47 did not cluster with any other sequences in the analysis. See the legend of Fig. 3 for explanation.

each by one to nine clones and the sequence similarities were 78–96%, when excluding the phylotype Rs-B62, which possessed an about 150-bp insertion in the 5′ region. All the α-proteobacterial phylotypes were found for the first time except for two, Rs-A44 and Rs-K73, which showed high sequence similarities to *Methylobacterium* sp. V3 (99.5%) and *Methylocella palustris* (97.2%), respectively (Fig. 11).

Two phylotypes represented by only one and two clones were affiliated with γ-Proteobacteria, and the phylotypes Rs-N33 and Rs-M74 were almost identical to the sequences from *Pseudomonas aeruginosa* (100%) and *Enterobacter amnigenus* (99.4%), respectively (Fig. 11). Among the four β-proteobacterial phylotypes, Rs-A34 was the most abundant from the sequenced clones (Fig. 4), and affiliated with the *Rhodocyclus* group (Fig. 11). In the *Rhodocyclus* group, the phylotype Rs-K54 shared 95.6% sequence identity with BCf6-08 (AB062831) from *C. formosanus*. Rs-D51 was found to be related to *Chromobacterium violaceum* (M22510) and almost identical to the sequence OS1L-24 (AB086875) (99.5%) from activated sludge. The three ε-proteobacterial phylotypes were represented each by one to five clones and formed a monophyletic cluster, showing sequence similarities of 93–95% (Fig. 11). There was no cultured bacterium showing a high sequence similarity to them. The δ-proteobacterial phylotypes were comprised of five *Desulfovibrio* and one related to *Geobacter* (Fig. 11). Among the *Desulfovibrio* phylotypes, Rs-N31 was the most abundant (Fig. 4), and is related to *Desulfovibrio intestinalis* isolated from the gut of the termite *M. darwiniensis* [34], although the sequence similarity was only 88%.

3.2.6. Actinobacteria

Fourteen phylotypes, represented by one to three clones, were affiliated with Actinobacteria (Fig. 12). Five phylotypes were clustered with *Atopobium minutum* (S44204) and BCf9-11 (AB062844) from *C. formosanus*. The other five phylotypes were clustered with RFS10 (AF068463), part of a chimeric sequence from the termite *R. flavipes* [9]. They exclusively constituted a robust monophyletic cluster with sequence similarities of more than 88% (Fig. 12). Rs-N96 was the closest to RFS10, sharing 95.9% sequence identity.

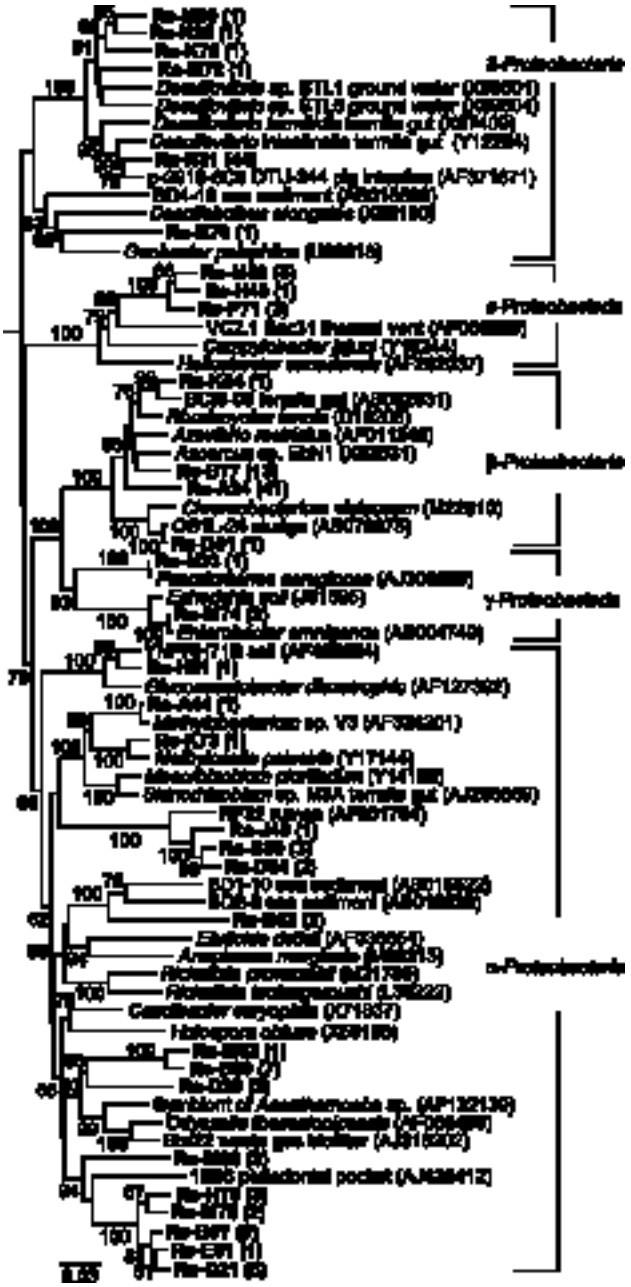


Fig. 11. Phylogenetic tree showing the positions of 16S rDNA phylotypes affiliated with Proteobacteria, recovered from the gut of the termite *R. speratus*. *Desulfovibrio intestinalis* [34] and *D. termitidis* were isolated from the termite *M. darwiniensis*. *Sinorhizobium* sp. M3A was isolated from the termite *Nasutitermes nigriceps*. See the legend of Fig. 3 for explanation.

3.2.7. Phylogenetic groups with few representative clones

The phylotype Rs-M39 was clustered with *Acidobacterium capsulatum* (D26171) and shared 97.0% sequence identity with the clone TRB82 (AF047646), recovered from an acid mine drainage site (Fig. 10). Rs-M47, represented by only one clone, did not cluster with any other sequences in the analysis (Fig. 10). The sequence showing the highest homology to Rs-M47 in the BLAST analysis was sequence BD4-10 (AB015569), recovered from deep sea sediment [35], although the similarity was as low as 79%.

Four phylotypes, affiliated with Planctomycetes, were clustered with BCf2-25 (AB062813) from *C. formosanus*, constituting a robust monophyletic cluster (Fig. 13A). The sequence similarities within this cluster were 89–96%. Two phylotypes were affiliated with Verrucomicrobia, and Rs-D37 was clustered with BCf3-05 (AB062814) from *C. formosanus*, with a high bootstrap confidence value (Fig. 13B), sharing 82.5% sequence identity. Two phylotypes were affiliated with the candidate division OP11 [36] (Fig. 13C). The phylotype Rs-J47 was clustered with the oral clone X112 (91% similarity), and Rs-M28 was clustered with clones BD1-5 [35] and VC1.2-cl13 [37], recovered from deep sea sediment (89% similarity). The phylotype Rs-H34 was clustered with *Gleobacter violaceus* in the phylum Cyanobacteria (Fig. 13D), and shared 94.7% sequence identity with the bovine ruminal clone 4COd-2 [27]. The remaining two phylotypes, Rs-A23 and Rs-J96, did not cluster with any known sequences with a high bootstrap confidence value (Fig. 13D). The sequences listed in the results of the BLAST and FASTA analysis of these two phylotypes were from Clostridia, including *Caloramator* sp. JW/JH-1 (AF286863), *Caloramator coolhaasii* (AF104215) and *Clostridium thermopalmarium* (X72869), while the sequence similarities were lower than 80%.

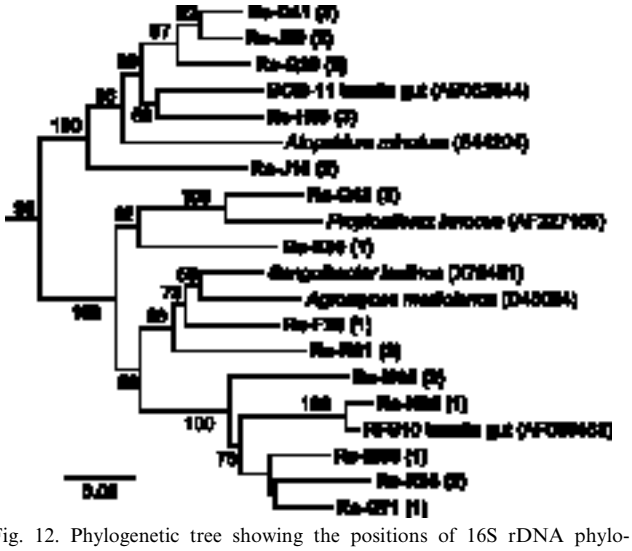


Fig. 12. Phylogenetic tree showing the positions of 16S rDNA phylotypes affiliated with Actinobacteria, recovered from the gut of the termite *R. speratus*. RFS10 is a short sequence (corresponding to positions 64–1111 in *E. coli* J01695), recovered as part of a chimeric sequence from the termite *R. flavipes* [9]. See the legend of Fig. 3 for explanation.

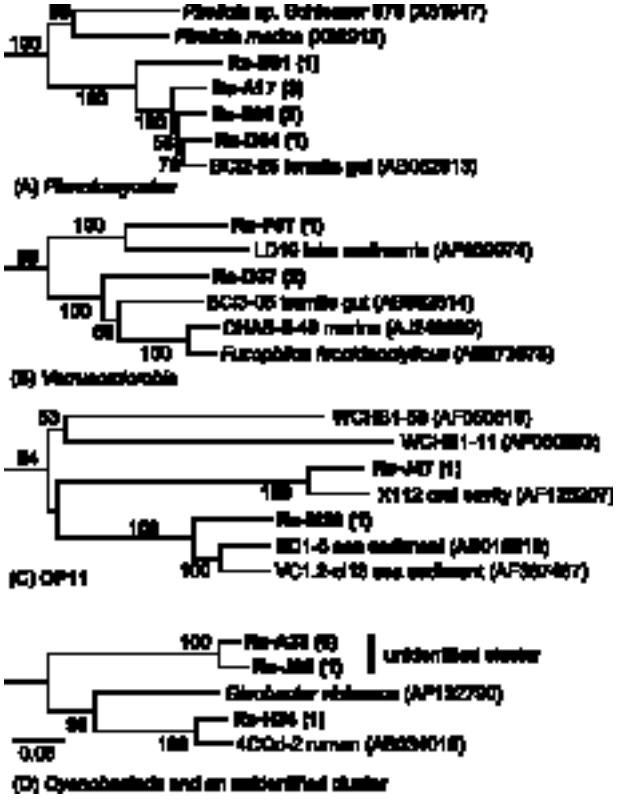


Fig. 13. Phylogenetic tree showing the positions of 16S rDNA phylotypes affiliated with (A) Planctomycetes, (B) Verrucomicrobia, (C) the candidate division OP11 and (D) Cyanobacteria, recovered from the gut of the termite *R. speratus*. The 'unidentified cluster' was not related to any known sequences with a high bootstrap confidence value. See the legend of Fig. 3 for explanation.

In total, 248 of the 268 phylotypes were found for the first time in the present study. Some of these phylotypes were clustered with each other and/or with sequences derived from other termite species. This implies the existence of termite-specific bacterial lineages. However, it is currently impossible to examine the coevolutionary relationship between termites and the bacterial symbionts because only small amounts of sequence data from other termite species are available. We expect that this study becomes the basis for further studies on the diversity of termite gut microbiota and the phylogenetic relationships of these bacteria, as well as the localization and characterization of the bacteria.

Acknowledgements

We thank Ms. Hiroe Yuzawa, Ms. Chutima Disyen, Ms. Khajarin Sirihongsuwan and Ms. Chanpen Boontong for assisting with sequencing. This work was partially supported by grants for the Bioarchitect Research Program and the Eco Molecular Science Research Program from RIKEN.

References

[1] Eutick, M.L., Veivers, P.C., O'Brien, R.W. and Slaytor, M. (1978) Dependence of the higher termite, *Nasutitermes exitiosus* and the lower termite, *Coptotermes lacteus* on their gut flora. J. Insect Physiol. 24, 363–368.

[2] Breznak, J.A. (2000) Ecology of prokaryotic microbes in the guts of wood- and litter-feeding termites. In Termites: Evolution, Sociality, Symbioses, Ecology (Abe, T., Bignell, D.E. and Higashi, M., Eds), pp. 209–232. Kluwer Academic, Dordrecht.

[3] Brune, A. and Friedrich, M. (2000) Microecology of the termite gut: structure and function on a microscale. Curr. Opin. Microbiol. 3, 263–269.

[4] Leadbetter, J.R., Schmidt, T.M., Graber, J.R. and Breznak, J.A. (1999) Acetogenesis from H_2 plus CO_2 by spirochetes from termite guts. Science 283, 686–689.

[5] Lilburn, T.G., Kim, K.S., Ostrom, N.E., Byzek, K.R., Leadbetter, J.R. and Breznak, J.A. (2001) Nitrogen fixation by symbiotic and free-living spirochetes. Science 292, 2495–2498.

[6] Ohkuma, M. and Kudo, T. (1996) Phylogenetic diversity of the intestinal bacterial community in the termite *Reticulitermes speratus*. Appl. Environ. Microbiol. 62, 461–468.

[7] Noda, S., Ohkuma, M., Usami, R., Horikoshi, K. and Kudo, T. (1999) Culture-independent characterization of a gene responsible for nitrogen fixation in the symbiotic microbial community in the gut of the termite *Neotermes koshunensis*. Appl. Environ. Microbiol. 65, 4935–4942.

[8] Berchtold, M., Chatzinotas, A., Schonhuber, W., Brune, A., Amann, R., Hahn, D. and Konig, H. (1999) Differential enumeration and in situ localization of microorganisms in the hindgut of the lower termite *Mastotermes darwiniensis* by hybridization with rRNA-targeted probes. Arch. Microbiol. 172, 407–416.

[9] Lilburn, T.G., Schmidt, T.M. and Breznak, J.A. (1999) Phylogenetic diversity of termite gut spirochaetes. Environ. Microbiol. 1, 331–345.

[10] Brauman, A., Dore, J., Eggleton, P., Bignell, D.E., Breznak, J.A. and Kane, M.D. (2001) Molecular phylogenetic profiling of prokaryotic communities in guts of termites with different feeding habits. FEMS Microbiol. Ecol. 35, 27–36.

[11] Iida, T., Ohkuma, M., Ohtoko, K. and Kudo, T. (2000) Symbiotic spirochetes in the termite hindgut: phylogenetic identification of ectosymbiotic spirochetes of oxymonad protists. FEMS Microbiol. Ecol. 34, 17–26.

[12] Maidak, B.L., Cole, J.R., Lilburn, T.G., Parker, C.T.J., Saxman, P.R., Farris, R.J., Garrity, G.M., Olsen, G.J., Schmidt, T.M. and Tiedje, J.M. (2001) The RDP-II (Ribosomal Database Project). Nucleic Acids Res. 29, 173–174.

[13] Altschul, S.F., Madssen, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

[14] Lipman, D.J. and Person, W.R. (1985) Rapid and sensitive protein similarity searches. Science 227, 1435–1441.

[15] Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

[16] Kumar, S., Tamura, K., Jakobson, I.B. and Nei, M. (2001) *MEGA2*: *Molecular Evolutionary Genetics Analysis* software. Bioinformatics 17, 1244–1245.

[17] Singleton, D.R., Furlong, M.A., Rathbun, S.L. and Whitman, W.B. (2001) Quantitative comparisons of 16S rRNA gene sequence libraries from environmental samples. Appl. Environ. Microbiol. 67, 4374–4376.

[18] Good, I.J. (1953) The population frequencies of species and the estimation of population parameters. Biometrika 40, 237–264.

[19] Paster, B.J., Dewhirst, F.E., Cooke, S.M., Fussing, V., Poulsen, L.K. and Breznak, J.A. (1996) Phylogeny of not-yet-cultured spirochetes from termite guts. Appl. Environ. Microbiol. 62, 347–352.

[20] Curtis, T.P., Sloan, W.T. and Scannell, J.W. (2002) Estimating prokaryotic diversity and its limits. Proc. Natl. Acad. Sci. USA 99, 10494–10499.

[21] Leser, T.D., Amenuvor, J.Z., Jensen, T.K., Lindecrona, R.H., Boye, M. and Moller, K. (2002) Culture-independent analysis of gut bacteria: the pig gastrointestinal tract microbiota revisited. Appl. Environ. Microbiol. 68, 673–690.

[22] Schultz, J.E. and Breznak, J.A. (1978) Heterotrophic bacteria present in hindguts of wood-eating termites [*Reticulitermes flavipes* (Kollar)]. Appl. Environ. Microbiol. 35, 930–936.

[23] Ward, B.B. (2002) How many species of prokaryotes are there? Proc. Natl. Acad. Sci. USA 99, 10234–10236.

[24] Ohkuma, M., Iida, T. and Kudo, T. (1999) Phylogenetic relationships of symbiotic spirochetes in the gut of diverse termites. FEMS Microbiol. Lett. 181, 123–129.

[25] Ohkuma, M., Noda, H., Hongoh, Y. and Kubo, T. (2002) Diverse bacteria related to the bacteroides subgroup of the CFB phylum within the gut symbiotic communities of various termites. Biosci. Biotechnol. Biochem. 66, 78–84.

[26] Tajima, K., Aminov, R.I., Nagamine, T., Ogata, K., Nakamura, M., Matsui, H. and Benno, Y. (1999) Rumen bacterial diversity as determined by sequence analysis of 16S rDNA libraries. FEMS Microbiol. Ecol. 29, 159–169.

[27] Tajima, K., Arai, S., Ogata, K., Nagamine, T., Matsui, H., Nakamura, M., Aminov, R.I. and Benno, Y. (2000) Rumen bacterial community transition during adaptation to high-grain diet. Anaerobe 6, 273–284.

[28] Zhu, X.Y., Zhong, T., Pandya, Y. and Joerger, R.D. (2002) 16S rRNA-based analysis of microbiota from the cecum of broiler chickens. Appl. Environ. Microbiol. 68, 124–137.

[29] Hold, G.L., Pryde, S.E., Russell, V.J., Furrie, E. and Flint, H.J. (2002) Assessment of microbial diversity in human colonic samples by 16S rDNA sequence analysis. FEMS Microbiol. Ecol. 39, 33–39.

[30] Frohlich, J. and Konig, H. (1999) Rapid isolation of single microbial cells from mixed natural and laboratory populations with the aid of a micromanipulator. Syst. Appl. Microbiol. 22, 249–257.

[31] Bond, P.L., Hugenholtz, P., Keller, J. and Blackall, L.L. (1995) Bacterial community structures of phosphate-removing and non-phosphate-removing activated sludges from sequencing batch reactors. Appl. Environ. Microbiol. 61, 1910–1916.

[32] Godon, J.J., Zumstein, E., Dabert, P., Habouzit, F. and Moletta, R. (1997) Molecular microbial diversity of an anaerobic digestor as determined by small-subunit rDNA sequence analysis. Appl. Environ. Microbiol. 63, 2802–2813.

[33] Dojka, M.A., Hugenholtz, P., Haack, S.K. and Pace, N.R. (1998) Microbial diversity in a hydrocarbon- and chlorinated-solvent-contaminated aquifer undergoing intrinsic bioremediation. Appl. Environ. Microbiol. 64, 3869–3877.

[34] Frohlich, J., Sass, H., Babenzien, H.D., Kuhnigk, T., Varma, A., Saxena, S., Nalepa, C., Pfeiffer, P. and Konig, H. (1999) Isolation of *Desulfovibrio intestinalis* sp. nov. from the hindgut of the lower termite *Mastotermes darwiniensis*. Can. J. Microbiol. 45, 145–152.

[35] Li, L., Kato, C. and Horikoshi, K. (1999) Bacterial diversity in deep-sea sediments from different depths. Biodivers. Conserv. 8, 659–677.

[36] Hugenholtz, P., Goebel, B.M. and Pace, N.R. (1998) Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. J. Bacteriol. 180, 4765–4774.

[37] Corre, E., Reysenbach, A.L. and Prieur, D. (2001) Epsilon-proteobacterial diversity from a deep-sea hydrothermal vent on the Mid-Atlantic Ridge. FEMS Microbiol. Lett. 205, 329–335.

[38] Suzuki, M.T. and Giovannoni, S.J. (1996) Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. Appl. Environ. Microbiol. 62, 625–630.

[39] Weidner, S., Arnold, W. and Puhler, A. (1996) Diversity of uncultured microorganisms associated with the seagrass *Halophila stipula-*

cea estimated by restriction fragment length polymorphism analysis of PCR-amplified 16S rRNA genes. Appl. Environ. Microbiol. 62, 766–771.

[40] Marchesi, J.R., Sato, T., Weightman, A.J., Martin, T.A., Fry, J.C., Hiom, S.J., Dymock, D. and Wade, W.G. (1998) Design and evaluation of useful bacterium-specific PCR primers that amplify genes coding for bacterial 16S rRNA. Appl. Environ. Microbiol. 64, 795–799.

[41] Ohkuma, M. and Kudo, T. (1998) Phylogenetic analysis of the symbiotic intestinal microflora of the termite *Cryptotermes domesticus*. FEMS Microbiol. Lett. 164, 389–395.

[42] Osborn, A.M., Moore, E.R. and Timmis, K.N. (2000) An evaluation of terminal-restriction fragment length polymorphism (T-RFLP) analysis for the study of microbial community structure and dynamics. Environ. Microbiol. 2, 39–50.

[43] Hethener, P., Brauman, A. and Garcia, J.L. (1992) *Clostridium termitidis* sp. nov., a cellulolytic bacterium from the gut of the wood-feeding Termite, *Nasutitermes lujae*. Syst. Appl. Microbiol. 15, 52–58.

[44] Bauer, S., Tholen, A., Overmann, J. and Brune, A. (2000) Characterization of abundance and diversity of lactic acid bacteria in the hindgut of wood- and soil-feeding termites by molecular and culture-dependent techniques. Arch. Microbiol. 173, 126–137.